UNIFORM BOUNDS FOR PERIOD INTEGRALS AND SPARSE EQUIDISTRIBUTION

JAMES TANIS AND PANKAJ VISHE

ABSTRACT. Let $M = \Gamma\backslash \mathrm{PSL}(2,\mathbb{R})$ be a compact manifold, and let $f \in \mathrm{C}^\infty(M)$ be a function of zero average. We use spectral methods to get uniform (i.e. independent of spectral gap) bounds for twisted averages of f along long horocycle orbit segments. We apply this to obtain an equidistribution result for sparse subsets of horocycles on M.

1. Introduction

Let $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$ be a cocompact lattice, and let $M = \Gamma\backslash \mathrm{PSL}(2,\mathbb{R})$. Let

$$n(t) := \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad a(t) := \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix},$$

denote the one-parameter subgroups generating the horocycle and the geodesic flows respectively.

Let $\mathrm{L}^2(M)$ be the space of complex-valued functions on M, which are square-integrable with respect to the $\mathrm{PSL}(2,\mathbb{R})$-invariant volume form. The space $\mathrm{L}^2(M)$ is a right regular representation of $\mathrm{PSL}(2,\mathbb{R})$ and any element of the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ of $\mathrm{PSL}(2,\mathbb{R})$ acts on $\mathrm{L}^2(M)$ as an essentially skew-adjiont differential operator. Let $\{Y, X, Z\}$ be a basis for the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ given by,

$$Y = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The center of the enveloping algebra for $\mathfrak{sl}(2,\mathbb{R})$ is generated by the Casimir element

$$\square := -Y^2 - 1/2(XZ + ZX)\,,$$

which acts by multiplication by a constant on each irreducible, unitary representation of $\mathrm{PSL}(2,\mathbb{R})$. These constants, $\mu \in \mathrm{spec}(\square) := \mathbb{R}_+ \cup \{(-n^2+2n)/4 : n \in \mathbb{Z}_+\}$ classify the nontrivial, unitary, irreducible representations of $\mathrm{PSL}(2,\mathbb{R})$ into three categories: A representation is called principal series if $\mu \geq 1/4$, complementary series if $0 < \mu < 1/4$, and discrete series if $\mu \leq 0$. The Casimir element takes the value zero on the trivial representation, which is spanned by the $\mathrm{PSL}(2,\mathbb{R})$-invariant volume form, denoted by dg.

Let Ω_Γ be the set of eigenvalues of $\square$ on $\mathrm{L}^2(M)$, counting multiplicities. Let

$$\mathrm{L}^2(M) = \mathbb{C} \bigoplus_{\mu \in \Omega_\Gamma} V_\mu, \tag{1.1}$$

Tanis is partially supported by the French ANR grant GeoDyM (ANR-11-BS01-0004). Vishe is supported by the EPSRC Programme Grant: EP/J018260/1.

be a Fourier decomposition of $L^2(M)$, where for each $0 < \mu$, V_μ is an irreducible, unitary representation of M in the μ-eigenspace of $\square$; and for $\mu \leq 0$, $V_\mu = V_\mu^+ \oplus V_\mu^-$ is a direct sum of two inequivalent irreducible unitary representations of M, called holomorphic and anti-holomorphic discrete series representations in the μ-eigenspace of $\square$.

We will use the L^p based norms for $p = 1, 2, \infty$. We write $S_{p,0}(f)$ for the L^p norm of f. Let $\mathscr{O}_s$ be the collection monomials in $\{Y, X, Z\}$ up to order $s \in \mathbb{Z}_{\geq 0}$. Let $W^{s,p}(M)$ be the space of functions with bounded norm

$$S_{p,s}(f) := \sum_{B \in \mathscr{O}_s} S_{p,0}(Bf)\,.$$

The Hilbert Sobolev spaces $W^{s,2}(M)$ are denoted by $W^s(M)$. For even integers s, they consist of functions f with bounded norms

$$S_{2,s}(f) = S_{2,0}\left((I - Y^2 - 1/2X^2 - 1/2Z^2)^{s/2} f\right)\,. \tag{1.2}$$

Using interpolation, these norms can be defined for all $s \geq 0$ (see [13]). [12, Lemma 6.3] implies that for all $s \in \mathbb{Z}_{\geq 0}$, there are constants $C_s, C_s' > 0$ such that

$$C_s S_{2,s}(f) \leq \sum_{B \in \mathscr{O}_s} S_{2,0}(Bf) \leq C_s' S_{2,s}(f)\,.$$

Moreover, $W^s(M)$ is endowed with an inner product, and by irreducibility, and (1.1), we have

$$W^s(M) = \mathbb{C} \bigoplus_{\mu \in \Omega_\Gamma} V_\mu^s\,, \tag{1.3}$$

where V_μ^s is the subspace $V_\mu \cap W^s(M)$.

The distributional dual space of $W^s(M)$ is denoted by $W^{-s}(M) := (W^s(M))'$, equipped with the natural distributional norm $S_{2,-s}$. Our distributions are defined and studied in $W^{-s}(M)$. The space of smooth functions on M is denoted by $C^\infty(M) := \cap_{s \geq 0} W^s(M)$, and its distributional dual space is $\mathscr{E}'(M) := \cup_{s \geq 0} W^{-s}(M)$. For each $\mu \in \Omega_\Gamma$, let $V_\mu^\infty := \bigcap_{s \geq 0} V_\mu^s$.

Our computations will be carried out in irreducible models consisting of functions defined on the real line. Using unitary equivalence, we use the same notation $S_{2,s}(f)$ for L^2-based norms in models. We start by listing our results.

1.1. **Bounds for period integrals.**

Let ψ be the additive character

$$\psi(t) := e^{iat}, \text{ for all } t \in \mathbb{R},$$

for some $a \in \mathbb{R} \setminus \{0\}$. For any $T \geq 1$ and for any $x \in M$, let $f \star \sigma_T(x)$ denote the unipotent period integral defined by

$$f \star \sigma_T(x) := \frac{1}{T} \int_0^T \psi(t) f(xn(t))dt\,. \tag{1.4}$$

Let $0 < \lambda_1 \leq 1/4$ be the spectral gap of the Laplacian on M, and let $\alpha_0 := \sqrt{1 - 4\lambda_1}$. Venkatesh [18, Lemma 3.1] used equidistribution and mixing of the horocycle flow to prove the following bound:

Lemma 1.1. *Let $f \in C^\infty(M)$ satisfying $\int_M f(g)\,dg = 0$. Then*

$$S_{\infty,0}(f \star \sigma_T) \ll_\Gamma T^{-b} S_{\infty,1}(f),$$

where $b < \frac{(1-2\alpha_0)^2}{8(3-2\alpha_0)}$, and the implied constant is independent of ψ.

Our main theorem is an estimate of the L^∞ norm of the derivatives of $f \star \sigma_T$:

Theorem 1.2. *Let $f \in C^\infty(M)$ be such that $\int_M f(g)\,dg = 0$. Then for any $k \in \mathbb{N}$, and for any $\varepsilon > 0$, we have*

$$S_{\infty,k}(f \star \sigma_T) \ll_{\Gamma,\varepsilon} (1 + |a|^{-1/2}) T^{2k-1/9+\varepsilon} S_{2,k+11/2+\varepsilon}(f). \tag{1.5}$$

Moreover, at the cost of a possibly larger, unspecified dependence on a, the above bound can be improved to get

$$S_{\infty,k}(f \star \sigma_T) \ll_{\Gamma,a,\varepsilon} T^{2k-(9-\sqrt{73})/4+\varepsilon} S_{2,k+11/2+\varepsilon}(f), \tag{1.6}$$

where $(9-\sqrt{73})/4 < 1/8.772$.

In particular, we remove the dependence of spectral gap in the exponent in Lemma 1.1, at the cost of a factor which depends on ψ. The dependence of Γ in the constant can be made explicit using the injectivity radius and the spectral gap. We remark that one should not expect an estimate independent of both the spectral gap and ψ, since, as $a \to 0$, the behavior of $f \star \sigma_T$ is increasingly governed by the rate of equidistribution of the horocycle flow on M, which depends on the spectral gap.

1.2. **Sparse equidistribution results.** Recently, there has been an increasing interest in studying equidistribution properties of sparse subsets of horocycle orbits. A question of Margulis [14] asks whether the sequence $\{x_0 n(t_j)\}_{j\in\mathbb{Z}^+}$ is dense in M, for

- t_j is the j_{th} prime number;
- $t_j = \lfloor j^{1+\gamma} \rfloor$, for some $\gamma > 0$.

A conjecture of Shah further predicts that the sequence $\{x_0 n(j^{1+\gamma})\}_{j\in\mathbb{Z}^+}$ is equidistributed for *any* $\gamma > 0$. Margulis' first question was partially answered by Sarnak-Ubis [16], by proving that the horocycle orbit of a non-periodic point at prime values is dense in a set of positive measure in the modular surface.

Margulis' second question was answered in [18, Theorem 3.1], where lemma 1.1 was used to achieve equidistribution of the points $\{x_0 n(j^{1+\gamma})\}$, for any $0 < \gamma < \gamma_{\max}(\Gamma)$, depending on the spectral gap of Γ. We use theorem 1.2 to remove the dependence of spectral gap in the above result, establishing equidistribution of points $\{x_0 n(j^{1+\gamma})\}$, for any $\gamma < 1/26$, giving further evidence for Shah's conjecture.

Theorem 1.3. *Let $b < \frac{(1-2\alpha_0)^2}{8(3-2\alpha_0)}$, and let $b_1 < 1/9$, then for any $x_0 \in M$, any $f \in C^\infty(M)$, and for any $0 \le \gamma < \frac{b+2b_1}{6-(b+2b_1)}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(x_0 n(j^{1+\gamma})) = \int_M f(g)\,dg.$$

In other words, the sequence $\{x_0 n(j^{1+\gamma}) : 0 \le j \le N\}$ is equidistributed in M as N tends to ∞.

It should be noted that the methods in this paper in fact give *effective* bounds for the rate of equidistribution in the above theorem, which depend on γ. However, we do not to mention it here in order to keep the statement of the theorem more accessible.

1.3. **Sparse equidistribution for smooth time-changes of horocycle flows.** Let $\tau : M \times \mathbb{R} \to \mathbb{R}$ be a smooth cocycle over $\{n(t)\}_{t\in\mathbb{R}}$, i.e. for all $x \in M$ and $t, s \in \mathbb{R}$,

$$\tau(x, t + s) = \tau(x, t) + \tau(xn(t), s) .$$

Assume that for all $x \in M$, $\tau(x, t)$ is a strictly increasing function of t. Let $\rho : M \to \mathbb{R}^+$ denote the positive function defined by

$$\rho(x) = \frac{d}{dt}\tau(x, t)_{|t=0} .$$

We assume $\rho \in W^6(M)$.

For all $x \in M$, the smooth time change $\{n^\rho_t\}_{t\in\mathbb{R}}$ of $\{n(t)\}$ is defined by

$$n^\rho_{\tau(x,t)}(x) := xn(t) .$$

We will also write $n^\rho(\tau) := n^\rho_\tau$. The vector field for $\{n^\rho_\tau\}_{\tau\in\mathbb{R}}$ is generated by

$$X_\rho := X/\rho ,$$

and the X_ρ-invariant volume form is $d_\rho g := \rho\, dg$. In the wake of Shah's conjecture, it is natural to further ask whether Shah's conjecture holds for any smooth time change of a horocycle flow.

Using the method in [18], and the results of Forni-Ulcigrai [7], we obtain a sparse equidistribution result for smooth time-changes of horocycle flow, dependent on the spectral gap, thus providing a partial answer to the above question.

Theorem 1.4. *Let $b = -(1 - \alpha_0)^2/200$. Then for any $x_0 \in M$, any $f \in C(M)$, and any $0 \le \gamma < b$,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} f(x_0 n^\rho(j^{1+\gamma})) = \int_M f(g)\, d_\rho g .$$

1.4. **Remarks.** The method used here is simple yet powerful, and could be employed in answering further questions related to the horocycle flows. For instance, proposition 3.2 below gives a bound for the mean-square estimate for twisted averages of the horocycle flow, improving [7, Theorem 4] in a very special case. Moreover, since analogous theory of Kirillov models is available for quotients of $\mathrm{PGL}(2, k)$, for a field k in a very general setting, these estimates are likely to be generalized there as well.

An independent work of the first author with L. Flaminio and G. Forni [6] also addresses the question of bounding period integrals and application to Shah's conjecture. With more work, [6] obtains stronger results by using a rescaling argument as in [5] for the *twisted horocycle flow*, namely, a combination of the horocycle flow with a circle translation on $\Gamma\backslash \mathrm{SL}(2, \mathbb{R}) \times S^1$.

Throughout the paper, we only deal with compact quotients of $\mathrm{PSL}(2, \mathbb{R})$. However, the non compact case can be dealt with analogously, using the corresponding spectral decomposition [4, Theorem 1.7]. In this case, certain period integrals correspond to the Fourier coefficients of automorphic forms (see [18, Section 3.2]). A non-compact version of theorem 1.2 would therefore provide uniform bounds for the Fourier coefficients of

automorphic forms. Even though the methods in this paper would fail to give better estimates than those of Good [10], and Bernstein and Reznikov [3]; their advantage lies in their simplicity, and general applicability.

It should be noted that in order to keep the exposition clearer, we have not tried to optimize the Sobolev norms appearing in the results in this paper. These can be improved using a more stringent approach.

1.5. **Acknowledgements.** The second author would like to thank A. Venkatesh for introducing him to the problem, for the discussions, and for the encouragement. The authors would like to thank G. Forni for the encouragement, and for his generous help. The authors are also grateful for M. Baruch for his help in understanding the Kirillov model for $\mathrm{PSL}(2,\mathbb{R})$.

2. Irreducible models and spectral decomposition

2.1. **Line models.** For a Casimir parameter $\mu > 0$, let $\nu = \sqrt{1-4\mu}$ be a representation parameter. The line model H_μ for a principal or complementary series representation space is realized on the Hilbert space consisting of functions on $\mathbb{R}$ with the following norm. If $\mu \geq 1/4$, then $\nu \in i\mathbb{R}$, and the corresponding norm is

$$S_{2,0}(f) = \|f\|_{L^2(\mathbb{R})}.$$

If $0 < \mu < 1/4$, then $0 < \nu < 1$, and

$$S_{2,0}(f) = \left(\int_{\mathbb{R}^2} \frac{f(x)\overline{f(y)}}{|x-y|^{1-\nu}}\, dx\, dy \right)^{1/2}.$$

The group action is defined by

$$\pi_\nu : \mathrm{PSL}(2,\mathbb{R}) \to \mathscr{U}(H_\mu)$$

$$\pi_\nu(A) f(x) = |-cx+a|^{-(\nu+1)} f\left(\frac{dx-b}{-cx+a}\right), \tag{2.1}$$

where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2,\mathbb{R})$, and $x \in \mathbb{R}$. Let H_μ^∞ be the space of all smooth vectors in H_μ.

In the discrete series case the situation is a little bit more complicated. For $4\mu = -n^2 + 2n$, where $n \in \mathbb{Z}_{\geq 2}$, let H_μ^∞ be the space of smooth functions f on $\mathbb{R}$, such that $x^{-n} f(-1/x)$ is smooth. The corresponding group action π_n is defined by

$$\pi_n(A) f(x) = |-cx+a|^{-n} f\left(\frac{dx-b}{-cx+a}\right). \tag{2.2}$$

H_μ^∞ consists of two irreducible invariant subspaces for the action π_n, denoted by $H_\mu^{+,\infty}$, and $H_\mu^{-,\infty}$. These representation spaces correspond to the smooth vectors in holomorphic and anti-holomorphic discrete series representations H_μ^+ and H_μ^-, for the eigenvalue $\mu = (-n^2+2n)/4$, respectively. $H_\mu^\pm$ can be shown to be unitarily equivalent to $V_\mu^\pm$. See [2, section 4], and [17] for more details.

2.2. **Kirillov Models.** The Kirillov model, denoted by K_μ, is closely related to the Fourier transform of the line model.

2.2.1. *Principal and complementary series representations.* For $\mu > 0$, we let

$$\phi : H_\mu \to K_\mu : f \to C_\mu |x|^{(1-\nu)/2} \hat{f}\,, \tag{2.3}$$

where C_μ is a constant which is defined to be 1 for $\mu \geq 1/4$, and it will be defined later for $0 < \mu < 1/4$. Then

$$K_\mu := \{\phi(f) : f \in H_\mu\}$$

with the norm

$$S_{2,0}(f) := \left(\int_{\mathbb{R}/\{0\}} |f(x)|^2 \frac{dx}{|x|} \right)^{1/2} .$$

We begin by showing that ϕ is unitary in the complementary series case, the principal series case being simpler. For $0 < \mu < 1/4$, let $R(x) = |x|^{\nu-1}$ be a homogeneous function on $\mathbb{R} \setminus \{0\}$. An easy computation shows that $\hat{R}(\xi) = |\xi|^{-\nu} \hat{R}(1)$. Moreover, $\hat{R}(1)$ is non zero since $\hat{R}$ is not identically zero. Then clearly

$$\|f\|_{H_\mu}^2 = \int_{\mathbb{R}^2} |x-y|^{\nu-1} f(x) \overline{f(y)}\, dx\, dy = \langle R * f, f \rangle\,,$$

where $\langle\ ,\ \rangle$ denotes the usual L^2 inner product on $\mathbb{R}$. The Plancherel theorem implies

$$\begin{aligned} \|f\|_{H_\mu}^2 &= \int_{\mathbb{R}\setminus\{0\}} |\hat{f}(\xi)|^2 \hat{R}(\xi)\, d\xi = \hat{R}(1) \int_{\mathbb{R}\setminus\{0\}} |\xi|^{-\nu} |\hat{f}(\xi)|^2\, d\xi \\ &= \hat{R}(1) \int_{\mathbb{R}\setminus\{0\}} \left| |\xi|^{(1-\nu)/2} \hat{f}(\xi) \right|^2 \frac{d\xi}{|\xi|} = \int_{\mathbb{R}\setminus\{0\}} |\phi(f)(\xi)|^2 \frac{d\xi}{|\xi|}\,, \end{aligned}$$

upon choosing $C_\mu = |\hat{R}(1)|^{1/2}$, proving that ϕ is unitary. The action of $g \in \mathrm{PSL}(2, \mathbb{R})$ on K_μ is given by

$$g \cdot \phi(f) := \phi(g \cdot f)\,,$$

implying that ϕ is an unitary equivalence. The explicit action of $n(t)$, and $a(t)$, on K_μ is given by

$$n(t) \cdot f(x) = e^{-itx} f(t), \qquad a(t) \cdot f(x) = f(e^t x)\,. \tag{2.4}$$

The explicit action of the basis X, Y, Z of $\mathfrak{sl}(2, \mathbb{R})$ on this model is given by:

$$X = -ix\,, \qquad Y = x\frac{\partial}{\partial x}\,, \qquad Z = i\frac{\mu}{x} - ix\frac{\partial^2}{\partial x^2}\,. \tag{2.5}$$

2.2.2. *Discrete series representation.* A detailed description of these models can be found in [14, Sections 4, 5]. We will merely state the various results here. As before, for $4\mu = -n^2 + 2n$, we let

$$\phi : H_\mu \to \mathrm{L}^2(\mathbb{R} \setminus \{0\}, dx/|x|) : f \to |x|^{1-n/2} \hat{f}\,.$$

[14, (4.8)] and [14, (4.11)] imply that ϕ maps $H_\mu^{+,\infty}$ into $\mathrm{L}^2((0, \infty), dx/|x|)$, and $H_\mu^{-,\infty}$ into $\mathrm{L}^2((-\infty, 0), dx/|x|)$. Upon completion, this gives us the following explicit description of the Kirillov model for V_μ:

The Kirillov model for a direct sum of holomorphic and anti-holomorphic discrete series representations of $\mathrm{PSL}(2, \mathbb{R})$, K_μ, is also realized on the space $\mathrm{L}^2(\mathbb{R} \setminus \{0\}, dx/|x|)$. The action of the Borel subgroup, and the Lie algebra here is analogous to (2.4), and (2.5).

It should be noted that the Kirillov model the homolorphic and anti-holomorphic discrete series representations, is realized via ϕ, on the spaces $\mathrm{L}^2((0,\infty), dx/|x|)$ and $\mathrm{L}^2((-\infty,0), dx/|x|)$ respectively.

2.2.3. *Bounds for the elements in the Kirillov model.* Throughout the paper, many of our computations in the space V_μ would be carried out in the Kirillov models K_μ, via unitary equivalence. As noted before, these models respect the L^2 based norms on V_μ. However, using Sobolev embedding on $\mathbb{R}$, we can get bounds for L^∞ norms on these models:

Lemma 2.1. *For any Kirillov model of* $\mathrm{PSL}(2,\mathbb{R})$,

$$\|f\|_{L^\infty(\mathbb{R})} \ll S_{2,1}(f)\,.$$

Proof. Let f be a smooth function in a Kirillov model for $\mathrm{PSL}(2,\mathbb{R})$. For any $x \in \mathbb{R}$, recall that $Xf(x) = -ixf(x)$, and $Yf(x) = xf'(x)$. Let $x_0 \in \mathbb{R}$ and let $x_1 := \min\{3|x_0|/4, 1/2\}$. Let $I_{x_0} := (x_0 - x_1, x_0 + x_1)$, and let h be in $C_c^\infty(I_{x_0})$ be such that $h(x_0) = 1$, and $\partial^k h(x) \ll |x_1|^{-k}$ for all $k \in \mathbb{Z}_{\geq 0}$.

We deal with the case $|x_0| < 1$ first. By Sobolev inequality,

$$\begin{aligned}
|f(x_0)| \leq \|fh\|_{\mathrm{L}^\infty} &\ll \|fh\|_{\mathrm{L}^1} + \|(fh)'\|_{\mathrm{L}^1} \\
&\ll \|f\|_{\mathrm{L}^1(I_{x_0})} + \|f'\|_{\mathrm{L}^1(I_{x_0})} + |x_0|^{-1}\|f\|_{\mathrm{L}^1(I_{x_0})} \\
&\ll \|f\|_{\mathrm{L}^1(I_{x_0})} + |x_0|^{-1}\|Yf\|_{\mathrm{L}^1(I_{x_0})} + |x_0|^{-1}\|f\|_{\mathrm{L}^1(I_{x_0})}\,.
\end{aligned}$$

We consider the term $\|f\|_{\mathrm{L}^1(I_{x_0})}$:

$$\begin{aligned}
\|f\|_{\mathrm{L}^1(I_{x_0})} &= \int_{x_0-x_1}^{x_0+x_1} |f(x)|dx \\
&\ll |x_0|^{1/2} \int_{x_0-x_1}^{x_0+x_1} |x|^{-1/2}|f(x)|dx \\
&\ll |x_0|^{1/2} \left(\int_{x_0-x_1}^{x_0+x_1} |x|^{-1}|f(x)|^2 dx\right)^{1/2} x_1^{1/2} \\
&\ll |x_0|\|f\|_{\mathrm{L}^2(I_{x_0}, dx/|x|)}\,.
\end{aligned}$$

After analogously bounding the rest of the terms, we get

$$\begin{aligned}
\|fh\|_{L^\infty} &\ll |x_0|\|f\|_{\mathrm{L}^2(I_{x_0},dx/|x|)} + \|Yf\|_{\mathrm{L}^2(I_{x_0},dx/|x|)} + \|f\|_{\mathrm{L}^2(I_{x_0},dx/|x|)} \\
&\ll S_{2,1}(f)\,.
\end{aligned}$$

When $|x_0| > 1$, the Sobolev embedding implies

$$\begin{aligned}
\|fh\|_{L^\infty} &\ll \|fh\|_{\mathrm{L}^1} + \|(fh)'\|_{\mathrm{L}^1} \ll \|f\|_{\mathrm{L}^1(I_{x_0})} + \|f'\|_{\mathrm{L}^1(I_{x_0})} \\
&\ll \int_{I_{x_0}} |xf(x)||x|^{-1}dx + \int_{I_{x_0}} |xf'_\mu(x)||x|^{-1}dx \\
&\ll S_{2,1}(f)\,,
\end{aligned}$$

proving the lemma. □

2.3. **Invariant distributions for the horocycle flow.** Let σ_{pp} be the set of non negative eigenvalues of the Laplace-Beltrami operator on M, which coincides with the set of non negative eigenvalues of the Casimir operator. The distributions invariant under the horocycle flow have been classified by Flaminio-Forni [4]. They showed that this space, $\mathscr{I}(M)$, has an infinite countable dimension, and that there is a decomposition

$$\mathscr{I}(M) = \bigoplus_{\mu\in\sigma_{pp}} \mathscr{I}_\mu \oplus \bigoplus_{n\in\mathbb{Z}^+} \mathscr{I}_n \,,$$

where

- for $\mu = 0$, the space $\mathscr{I}_0$ is spanned by scalar multiples of $\mathrm{PSL}(2,\mathbb{R})$-invariant volume, denoted by vol;
- for $0 < \mu < 1/4$, there is a splitting $\mathscr{I}_\mu = \mathscr{I}_\mu^+ \oplus \mathscr{I}_\mu^-$, where $\mathscr{I}_\mu^\pm \subset W^{-s}(M)$ if and only if $s > \frac{1\pm\sqrt{1-4\mu}}{2}$, and each subspace has dimension equal to the multiplicity of $\mu\in\sigma_{pp}$;
- for $\mu \geq 1/4$, the space $\mathscr{I}_\mu \subset W^{-s}(M)$ if and only if $s > 1/2$, and it has dimension equal to twice the multiplicity of $\mu\in\sigma_{pp}$;
- for $n\in\mathbb{Z}_{\geq 2}$, the space $\mathscr{I}_n \subset W^{-s}(M)$ if and only if $s > n/2$ and it has dimension equal to twice the rank of the space of holomorphic sections of the $n_{\rm th}$ power of the canonical line bundle over M.

For $s > 1/2$, let $\mathscr{I}_\mu^s := \mathscr{I}_\mu \cap W^{-s}(M)$ and $\mathscr{I}_n^s := \mathscr{I}_n \cap W^{-s}(M)$. By [4, Theorem 1.4], for all $\mu \neq \frac{1}{4}$, and $n\in\mathbb{Z}_{\geq 2}$, the spaces $\mathscr{I}_\mu^s$ and $\mathscr{I}_n^s$ have a basis of unit-normed (in $W^{-s}(M)$) eigenvectors for $\{a(t)\}_{t\in\mathbb{R}}$, which we denote by $\mathscr{B}_\mu^s$ and $\mathscr{B}_n^s$, respectively. The space $\mathscr{I}_{1/4}^s$ decomposes as $\mathscr{I}_{1/4}^{s,+}\cup\mathscr{I}_{1/4}^{s,-}$, where $\mathscr{I}_{1/4}^{s,-}$ has a basis of unit-normed eigenvectors for $\{a(t)\}_{t\in\mathbb{R}}$ denoted by $\mathscr{B}_{1/4}^{s,-}$, and $\mathscr{I}_{1/4}^{s,+}$ has a basis of unit-normed generalized eigenvectors, denoted by $\mathscr{B}_{1/4}^{s,+}$. Let

$$\mathscr{B}_{0,+}^s := \bigcup_{\mu\in\sigma_{pp}} B_\mu^s \cup \{\mathscr{B}_n^s : n = 2\} \,,$$

and

$$\mathscr{B}_+^s := \bigcup_{\mu\in\sigma_{pp}/\{0\}} \mathscr{B}_\mu^s \,.$$

2.4. **Spectral decomposition for averages of horocycle flow.** Let x_0 be a fixed arbitrary point of M. Then for any $T\geq 1$, let ν_T be defined on $L^2(M)$ by

$$\nu_T(f) := \frac{1}{T}\int_0^T f(x_0 n(t))\,dt \,.$$

For $\mu\in\sigma_{pp}$, and for $\mathscr{D}\in\mathscr{I}_\mu^\pm$, let $S_{\mathscr{D}} := \frac{1\pm\mathrm{Re}\sqrt{1-4\mu}}{2}$.

For any $s > 2$, we may project ν_T orthogonally in $W^{-s}(M)$ onto the basis $\mathscr{B}^s(M)$ and the orthogonal complement of its closed linear span, $\mathscr{I}^s(M)^\perp$. Then for all $\mu\in\Omega_\Gamma$, there exists distributions

$$\mathscr{D} := \begin{cases} \mathscr{D}_{x_0,T,\mu}^\pm \in \mathscr{I}_\mu^\pm, & \text{if } \mu\in\sigma_{pp}\,, \\ \mathscr{D}_{x_0,T,\mu} \in \mathscr{I}_n, & \text{if } 4\mu = -n^2+2n \text{ for } n\in\mathbb{Z}_{\geq 2}\,, \end{cases}$$

and $\mathscr{R}^s_{x_0,T} \in W^{-s}(M)$ is such that

$$\nu_T = \left(\mathrm{vol} + \sum_{\mathscr{D}} \mathscr{D}\right) \oplus \frac{\mathscr{R}^s_{x_0,T}}{T} \tag{2.6}$$

in the $W^{-s}(M)$ Sobolev structure.

[4] showed that for any $s' > 3$, $S_{2,-s'}(\mathscr{R}^{s'}(x_0,T)) \ll_s 1$. We now use arguments in [4, Section 5] and [5, Lemma 3.7], to prove that $\mathscr{R}^{s'}(x_0,T) \in W^{-s'}(M)$, for any $s' > 2$, along with a suitable bound for this norm. Being able to estimate the $W^{2+\varepsilon}$ norm of remainder distribution would enable us to get a stronger decay estimate in theorem 1.2.

Henceforth we assume $2 < s \leq 3$, and let $s_{\mathrm{reg}} := \sup_{\varepsilon>0}\{\lfloor 2s - \varepsilon \rfloor\}$. For $n \in \mathbb{Z}_{s_{\mathrm{reg}}}$, the above description of invariant distributions shows $\mathscr{I}_n(M) \not\subset W^{-s}(M)$, implying that these distributions do not appear in the decomposition (2.6) of ν_T in $W^{-s}(M)$. Therefore, using the definition of $\mathscr{R}^s_{x_0,T}$, we get

$$\mathscr{R}^s_{x_0,T} \in \bigoplus_{n=3}^{s_{\mathrm{reg}}} \mathscr{I}_n(M) \oplus \mathscr{I}^s(M)^{\perp}\,. \tag{2.7}$$

Lemma 2.2. *Let $2 < s \leq 3$. Let $\mathscr{R}^s_{x_0,T}$ be as in (2.6). Then for all $f \in C^\infty(M)$,*

$$|\mathscr{R}^s_{x_0,T}(f)| \ll_s \frac{1}{\sqrt{1-\sqrt{1-\lambda_1}}} S_{2,s}(f)\,.$$

Proof. We begin by observing that for any $n \in \mathbb{Z}_{\geq 2}$, for any $s > n/2$, and for any $f \in W^s(M)$, $\mathscr{R}^s_{x_0,T}\,|_{\mathscr{I}^s_n}(f) = T\nu_T\,|_{\mathscr{I}^s_n}(f)$. [4, Lemma 5.12] further implies that

$$|T\nu_T\,|_{\mathscr{I}^s_n}(f)| \ll_s S_{2,s}(f).$$

These bounds clearly suffice for any $2 < s \leq 3$, and for any $n = 3, 4, 5$.

It is therefore enough to consider f to be a function on which each $\mathscr{D} \in \mathscr{I}^s(M)$ vanishes. For such a function f, $\int_0^T f(x_0 n(t))dt = \mathscr{R}^s_{x_0,T}(f)$. For any $s > 2$, [4, Theorem 4.1] imples the existence of a function $g \in C^\infty(M)$ (unique up to additive constants) satisfying

$$Xg = f, \tag{2.8}$$

such that for any $0 \leq t < s - 1$,

$$S_{2,t}(g) \ll_{t,s} \frac{1}{\sqrt{1-\sqrt{1-\lambda_1}}} S_{2,s}(f)\,. \tag{2.9}$$

The fundamental theorem of calculus then implies

$$|\mathscr{R}^s_{x_0,T}(f)| = |\int_0^T Xg(x_0 n(t))dt| = |g(x_0 n(T)) - g(x_0)|.$$

Now we estimate each term on the right-hand side. Let $\tau_0 \in \{0, T\}$, and let $x_{\tau_0} := x_0 n(\tau_0)$. As in the proof of [5, Lemma 3.7], the mean value theorem implies that

$$\int_0^1 g(x_{\tau_0} n(\tau))d\tau = g(x_{\tau_0} n(\tau_1))\,,$$

for some $\tau_1 \in (0,1)$. A further application of the fundamental theorem of calculus, and (2.8) gives us

$$\int_0^1 g(x_{\tau_0} n(\tau))\, d\tau + \int_{\tau_1}^0 f(x_{\tau_0} n(\tau))\, d\tau = g(x_{\tau_0} n(\tau_1)) + \int_{\tau_1}^0 Xg(x_{\tau_0} n(\tau)) dt = g(x_{\tau_0}) \tag{2.10}$$

We now apply the Sobolev trace theorem in [5, Lemma 3.7] to the operator which maps g to the trace $\int_0^1 g(x_{t_0} n(t))dt$, to get that for any $\varepsilon > 0$,

$$|\int_0^1 g(x_{\tau_0} n(t))\, dt| \ll_\varepsilon S_{2,0}\left((I - Z^2 - Y^2)^{1/2+\varepsilon} g\right) \ll_\varepsilon S_{2,1+\varepsilon}(g) \ll_{\varepsilon,\lambda_1} S_{2,2+2\varepsilon}(f). \tag{2.11}$$

The trace theorem also analogously implies that

$$|\int_{\tau_1}^0 f(x_{\tau_0} n(\tau))\, d\tau| \ll_{\varepsilon,\lambda_1} S_{2,2+2\varepsilon}(f)\,.$$

Combining these bounds, we get $|g(x_{\tau_0})| \ll_{\varepsilon,\lambda_1} S_{2,2+2\varepsilon}(f)$, thus implying

$$|\mathscr{R}^s_{x_0,T}(f)| \ll_{\varepsilon,\lambda_1} S_{2,2+2\varepsilon}(f)\,,$$

where the implied dependence on λ_1 is $C_{\lambda_1} = \frac{1}{\sqrt{1-\sqrt{1-\lambda_1}}}$, thus proving the lemma for any $s > 2$, upon suitably choosing ε. □

2.4.1. *Explicit spectral decomposition for ν_T.* In light of the improved regularity of the remainder distribution in lemma 2.2, [4, Theorem 1.4] implies that for any $s > 2$ and $(x,T) \in M \times \mathbb{R}_+$, there exists distributions $\mathscr{D}^s_{2,x_0,T} \in \mathscr{B}^s_{0,+}/\mathscr{B}^s_+$ and $\mathscr{R}^s_{x_0,T} \in W^{-s}(M)$ and a sequence of real-valued functions $\{c^s_{\mathscr{D}}(\cdot,\cdot)\}_{\mathscr{D}\in\mathscr{B}^s}$ on $M \times \mathbb{R}_{\geq 1}$ such that for all $f \in C^\infty(M)$,

$$\nu_T(f) = \int_M f(g)\, dg + \sum_{\mathscr{D}\in\mathscr{B}^s_+/\mathscr{B}^{s,+}_{1/4}} c^s_{\mathscr{D}}(x_0,T)\mathscr{D}(f)T^{-S_{\mathscr{D}}} + \sum_{\mathscr{D}\in\mathscr{B}^{s,+}_{1/4}} c^s_{\mathscr{D}}(x_0,T)\mathscr{D}(f)T^{-\frac{1}{2}}\log T$$
$$+ \frac{\mathscr{D}^s_{2,x_0,T}(f)\log T + \mathscr{R}^s_{x_0,T}(f)}{T}\,, \tag{2.12}$$

where,

$$\sum_{\mathscr{D}\in\mathscr{B}^s_+} |c^s_{\mathscr{D}}(x_0,T)|^2 + S_{2,-s}(\mathscr{D}^s_2(x_0,T)) + \ll_s 1\,, \quad S_{2,-s}(\mathscr{R}^s_{x_0,T}) \ll_{\lambda_1,s} 1$$

using [4, Corollary 5.3] and lemma 2.2.

Note that for any $0 < \mu < 1/4$, and for any $\mathscr{D} \in \mathscr{I}^{s,\pm}_\mu \cap \mathscr{B}^s_+$, the corresponding value of $S_{\mathscr{D}} = (1 \pm \nu)/2$. Thus the contribution in (2.12) from terms corresponding to $\mathscr{D} \in \mathscr{I}^{s,+}_\mu \cap \mathscr{B}^s_+$ is at most $O(T^{-1/2})$, for any $s > 1$. To summarize, for any $\varepsilon > 0$ we have

$$\nu_T(f) = \int_M f dg + \sum_{\mu\in\Omega_\Gamma\cap(0,1/4)} c^{2+\varepsilon}_{\mathscr{D}^-_\mu}(x_0,T)\mathscr{D}^-_\mu(f)T^{(-1+\nu)/2} + O_s(S_{2,1+\varepsilon}(f)T^{-\frac{1}{2}}\log^+ T)$$
$$+ O_s(S_{2,2+\varepsilon}(f)T^{-1}\log^+ T)\,, \tag{2.13}$$

where for every $\mu \in \Omega_\Gamma \cap (0,1/4)$, $\mathscr{D}^-_\mu \in \mathscr{I}^-_\mu$, $S_{2,-1-\varepsilon}(\mathscr{D}^-_\mu) \ll_s 1$, and $|c^{2+\varepsilon}_{\mathscr{D}^-_\mu}(x_0,T)| \ll_s 1$.

3. Proof of Theorem 1.2

We start by recalling that for any $f \in C^\infty(M)$,

$$f \star \sigma_T(x) = \frac{1}{T}\int_0^T \psi(t) f(xn(t))dt\,.$$

For any $f \in C^\infty(M)$, the following can be easily verified:

$$\begin{aligned} X(n(t)\cdot f) &= n(t)\cdot (Xf) \\ Y(n(t)\cdot f) &= n(t)\cdot ((Y+tX)f) \\ Z(n(t)\cdot f) &= n(t)\cdot ((Z-2tY-t^2X)f). \end{aligned} \tag{3.1}$$

These bounds imply that for any $s \in \mathbb{Z}_+$, any degree s monomial B_0 in X, Y, Z, and any $x \in M$,

$$|B_0(f\star\sigma_T)(x)| \ll \sum_{j=0}^{s}\sum_{B\in\mathscr{O}_j}\sum_{k=0}^{2j} |Bf\star\sigma_T^k(x)|, \tag{3.2}$$

where

$$f\star\sigma_T^k(x) := \frac{1}{T}\int_0^T \psi(t)t^k f(xn(t))dt\,. \tag{3.3}$$

Note that $f\star\sigma_T^0$ is equal to $f\star\sigma_T$. It is therefore enough to obtain a suitable bound for $S_{\infty,0}(Bf\star\sigma_H^k)$, for $k\in\mathbb{Z}_+$. Let $x_0\in M$ be a fixed arbitrary point. We begin by noting

$$\begin{aligned} &((f\star\sigma_H)\star\sigma_T^k)(x_0) \\ &\quad = \frac{1}{TH}\int_0^T\int_0^H \psi(t+z)t^k f(x_0n(t+z))\,dz\,dt \\ &\quad = \frac{1}{TH}\int_0^{T+H}\int_0^{\min\{y,H\}} \psi(y)(y-z)^k f(x_0n(y))\,dz\,dy \\ &\quad = \frac{1}{TH}\int_H^T\int_0^H \psi(y)y^k f(x_0n(y))\,dz\,dy + O(HT^{k-1}S_{\infty,0}(f)) \\ &\quad = f\star\sigma_T^k(x_0) + O(HT^{k-1}S_{\infty,0}(f)), \end{aligned}$$

implying $S_{\infty,0}(f\star\sigma_T^k - (f\star\sigma_H)\star\sigma_T^k) \ll HT^{k-1}S_{\infty,0}(f)$. An application of the Cauchy-Schwarz inequality implies

$$|((f\star\sigma_H)\star\sigma_T^k)(x_0)|^2 \ll T^{2k}\nu_T(|f\star\sigma_H|^2)\,. \tag{3.4}$$

In the light of (3.2), this implies that for any $B_0\in\mathscr{O}_s$,

$$|B_0(f\star\sigma_T)(x_0)| \ll \sum_{j=0}^{s}\sum_{B\in\mathscr{O}_j} \nu_T(|Bf\star\sigma_H|^2)^{1/2}T^{2j} + HT^{2s-1}S_{\infty,s}(f). \tag{3.5}$$

In order to bound the term $\nu_T(|Bf\star\sigma_H|^2)$, it is enough to get an appropriate bound for $\nu_T(|f\star\sigma_H|^2)$. We begin by noting that for $f\in \mathrm{L}^2(M)$, we have $|f\star\sigma_H|^2\in C^\infty(M)$.

We use the spectral decomposition in (2.13) to get that for any $\varepsilon > 0$,

$$\begin{aligned}\nu_T(|f \star \sigma_H|^2) = \int_M |f \star \sigma_H|^2 dg \\ + \sum_{\mu \in \Omega_\Gamma \cap (0,1/4)} c_{\mathscr{D}_\mu^-}(x_0, T) \mathscr{D}_\mu^-(|f \star \sigma_H|^2) T^{(-1+\nu)/2} + S_{2,1+\varepsilon}(|f \star \sigma_H|^2) T^{-\frac{1}{2}} \log^+ T \\ + S_{2,2+\varepsilon}(|f \star \sigma_H|^2) T^{-1} \log^+ T\,,\end{aligned} \tag{3.6}$$

using the Fourier expansion (1.1)

$$|f \star \sigma_H|^2 = \bigoplus_{\mu \in \Omega_\Gamma} (|f \star \sigma_H|^2)_\mu\,.$$

Lemma 3.1. *For any $\varepsilon > 0$, we have*

$$\begin{aligned}|\nu_T(|f \star \sigma_H|^2)| \ll_\varepsilon \int_M |f \star \sigma_H|^2 dg + T^{-\alpha_0} \sum_{\mu \in (0,1/4) \cap \Omega_\Gamma} c_\mu(x_0, T) \mathscr{D}_\mu^-(|f \star \sigma_H|^2) \\ + S_{2,1+\varepsilon}(|f \star \sigma_H|^2) T^{-\frac{1}{2}} \log^+ T + S_{2,2+\varepsilon}(|f \star \sigma_H|^2) T^{-1} \log^+(T)\,,\end{aligned}$$

where $\sum_\mu |c_\mu(x_0, T)|^2 \ll_\varepsilon 1$.

It remains to estimate the terms on the right-hand side of lemma 3.1.

3.1. **Estimate of $\int_M |f \star \sigma_H|^2\, dg$.** In order to bound $\int_M |f \star \sigma_H|^2\, dg = S_{2,0}(f \star \sigma_H)^2$, we start by obtaining a slightly more general bound $S_{2,s}(f \star \sigma_H)^2$, the utility of which will be evident in the later part of this section. Using the explicit action of the Lie algebra (3.1), a bound similar to (3.2) can be obtained for any $s \in \mathbb{N}$:

$$S_{2,s}(f \star \sigma_H)^2 \ll \sum_{j=0}^{s} \sum_{B \in \mathscr{O}_j} \sum_{k=0}^{2j} S_{2,0}(Bf \star \sigma_H^k)^2. \tag{3.7}$$

For any $f \in L^2(M)$, and any $\mu \in \sigma_{pp} \setminus \{0\}$, let f_μ be the projection of f in the component V_μ. Using the fact that the operators $\star \sigma_H^k$ are limits of discrete sums of operators of type $t^k n(t) \cdot f$, which map V_μ into V_μ, for any $\mu \in \Omega_\Gamma$, we can easily see that the operation $\star \sigma_H^k$ splits across irreducible components, i.e.

$$(f \star \sigma_H^k)_\mu = f_\mu \star \sigma_H^k. \tag{3.8}$$

Proposition 3.2. *Let $f \in C^\infty(M)$ be such that $\int f(g)\, dg = 0$. Then for any $s \in \mathbb{R}_+$,*

$$S_{2,s}(f \star \sigma_H)^2 \ll (1 + |a|^{-1}) H^{4s-1} S_{2,s+1}(f)^2\,.$$

Proof. Write

$$f = \bigoplus_{\mu \in \Omega_\Gamma} f_\mu\,,$$

where $f_\mu \in V_\mu$ for all $\mu \in \Omega_\Gamma$. By (3.8), we have

$$f \star \sigma_H^k = \bigoplus_{\mu \in \Omega_\Gamma} (f_\mu \star \sigma_H^k)\,.$$

In the light of (3.7), it is enough to get the corresponding bound for $S_{2,0}(f\star\sigma_H^k)^2$. Parseval's identity implies that,

$$S_{2,0}(f\star\sigma_H^k)^2=\sum_{\mu\in\Omega_\Gamma}S_{2,0}(f_\mu\star\sigma_H^k)^2\,.$$

For any $\mu\in\Omega_\Gamma$, we start estimating $S_{2,0}(f_\mu\star\sigma_H^k)^2$ in the Kirillov model for V_μ. For $f_\mu\in K_\mu$, the explicit action of $\star\,\sigma_H^k$ is given by

$$f\star\sigma_H^k(x)=\frac{1}{H}\int_0^H t^k e^{i(a-x)t}\,dt\,f(x). \tag{3.9}$$

Let $C\geq 3$, then if $|Ha|\leq C$, then we obtain the trivial bound

$$\int_{\mathbb{R}/\{0\}}|f_\mu\star\sigma_H^k(x)|^2\frac{dx}{|x|}\ll H^{2k}S_{2,0}(f_\mu)^2.$$

Henceforth, we assume $|Ha|\geq C$. Let $I_1=(a-1/H,a+1/H)$, $I_2=(-1/H,1/H)\setminus\{0\}$, and $I_3=\mathbb{R}\setminus(I_1\cup I_2\cup\{0\})$. Then the trivial bound

$$\int_{I_1}\left|\frac{1}{H}\int_0^H t^k e^{i(a-x)t}\,dt\,f_\mu(x)\right|^2\frac{dx}{|x|}\ll H^{2k-1}\|f_\mu\|^2_{\mathrm{L}^\infty(\mathbb{R})}(1+a^{-1}),$$

is enough upon further using lemma 2.1. On $I_2\cup I_3$, using repeated integration by parts, we get

$$\begin{aligned}
&\int_{I_2\cup I_3}\left|\frac{1}{H}\int_0^H t^k e^{i(a-x)t}\,dt\,f_\mu(x)\right|^2\frac{dx}{|x|}\\
&\ll\int_{\mathrm{I}_2\cup I_3}\left|\frac{1-e^{-i(a-x)H}}{iH(a-x)^{k+1}}f_\mu(x)\right|^2\frac{dx}{|x|}+\sum_{j=1}^{k}H^{2j}\int_{I_2\cup I_3}\left|\frac{f_\mu(x)}{H(x-a)^{k+1-j}}\right|^2\frac{dx}{|x|}\\
&\ll H^{2k-2}\int_{\mathrm{I}_2\cup I_3}\frac{|f_\mu|^2\,dx}{|x(x-a)^2|}\\
&\ll H^{2k-1}(1+a^{-1})\int_{\mathrm{I}_2}\frac{|f_\mu|^2\,dx}{|x|}+H^{2k-2}(1+a^{-1})\|f_\mu\|^2_{\mathrm{L}^\infty(\mathbb{R})}\int_{I_3}(|x|^{-2}+|x-a|^{-2})\,dx\\
&\ll H^{2k-1}(1+a^{-1})S_{2,0}(f_\mu)^2+(1+a^{-1})H^{2k-1}S_{2,1}(f_\mu)^2.
\end{aligned}$$

Combining these bounds, we get

$$S_{2,0}(f_\mu\star\sigma_H^k)^2\ll(1+|a|^{-1})H^{2k-1}S_{2,1}(f)^2.$$

Using these bounds, along with (3.7), we get that for any $s\in\mathbb{N}$,

$$S_{2,s}(f_\mu\star\sigma_H)^2\ll(1+|a|^{-1})H^{4s-1}S_{2,s+1}(f)^2.$$

Upon interpolation, we prove the above bound for any $s\in\mathbb{R}_+$, and upon further adding over all $\mu\in\Omega_\Gamma$ we get the proposition. $\square$

3.2. **Estimate of $\mathscr{D}_\mu^-(|f\star\sigma_H|^2)$.** In this subsection, we will estimate $\mathscr{D}_\mu^-(|f\star\sigma_H|^2)$ for $0<\mu<1/4$. Recall that $\nu=\sqrt{1-4\mu}$. Our main goal will be to establish the following proposition:

Proposition 3.3. *Let $f\in \mathrm{C}^\infty(\Gamma\backslash G)$ be a function of zero average, and let $0<\mu<1/4$. Then*

$$\sum_{\mu\in(0,1/4)\cap\Omega_\Gamma} |\mathscr{D}_\mu^-(|f\star\sigma_H|^2)| \ll_\Gamma (1+|a|^{-1})H^{-1}S_{2,(7-\nu)/2}(f)^2 .$$

We use the spectral decomposition to get

$$\mathscr{D}_\mu^-(|f\star\sigma_H|^2)=\sum_{\beta_1,\beta_2\in\Omega_\Gamma}\mathscr{D}_\mu^-(f_{\beta_1}\star\sigma_H\overline{f_{\beta_2}\star\sigma_H}).$$

In order to estimate $\mathscr{D}_\mu^-(f_{\beta_1}\star\sigma_H\overline{f_{\beta_2}\star\sigma_H})$, we start by defining a bi-sesquilinear functional $\mathscr{D}^-_{\mu,\beta_1,\beta_2}$ on $V_{\beta_1}^\infty\times V_{\beta_2}^\infty$, given by:

$$\mathscr{D}^-_{\mu,\beta_1,\beta_2}(f_1,f_2)=\mathscr{D}_\mu^-(f_1\bar{f}_2).$$

For any $b\in B$, $\mathscr{D}^-_{\mu,\beta_1,\beta_2}$ satisfies

$$\mathscr{D}^-_{\mu,\beta_1,\beta_2}((b\cdot f_1),(b\cdot f_2))=\chi_\mu(b)\mathscr{D}^-_{\mu,\beta_1,\beta_2}(f_1,f_2)$$

where

$$\chi_\mu(n(x)a(t))=e^{(-1+\nu)t/2}.$$

Let $\mathscr{E}_{\mu,\beta_1,\beta_2}$ be the space of bi-sesquilinear functionals $\mathscr{D}$ on $V_{\beta_1}^\infty\times V_{\beta_2}^\infty$ satisfying

$$\lambda\mathscr{D}(f,g)=\mathscr{D}(\lambda f,g)=\mathscr{D}(f,\overline{\lambda}g),\ \mathscr{D}(b\cdot f,b\cdot g)=\chi_\mu(b)\mathscr{D}(f,g), \tag{3.10}$$

for any b in the Borel subgroup of $\mathrm{PSL}(2,\mathbb{R})$, and $\lambda\in\mathbb{C}$. We begin by finding the dimension of $\mathscr{E}_{\mu,\beta_1,\beta_2}$:

Lemma 3.4. *For $\beta_1,\beta_2\in\Omega_\Gamma$, $\mathscr{E}_{\mu,\beta_1,\beta_2}$ is a two dimensional space. Moreover, if for $j=1,2$, $\phi_j:V_{\beta_j}\to K_{\beta_j}$ is the equivalence map, then the space $\mathscr{E}_{\mu,\beta_1,\beta_2}$ is spanned by the following two linearly independent functionals:*

$$B^1_{\mu,\beta_1,\beta_2}(f_{\beta_1},f_{\beta_2}):=\int_0^\infty |x|^{-(1+\nu)/2}(\phi_1 f_{\beta_1})(x)\overline{(\phi_2 f_{\beta_2})(x)}dx, \tag{3.11}$$

$$B^2_{\mu,\beta_1,\beta_2}(f_{\beta_1},f_{\beta_2}):=\int_{-\infty}^0 |x|^{-(1+\nu)/2}(\phi_1 f_{\beta_1})(x)\overline{(\phi_2 f_{\beta_2})(x)}dx. \tag{3.12}$$

Proof. We start by considering the space $\mathscr{F}_{\mu,\beta_1,\beta_2}$ of bi-sesquilinear functionals on the line models $H_{\beta_1}^\infty\times H_{\beta_2}^\infty$ satisfying (3.10), where $\nu_i=\sqrt{1-4\beta_i}$, if $\beta_i>0$, and $\nu_i=n_i-1$ for $n_i\in\mathbb{Z}^+$, if $\beta_i=-n_i^2+2n_i$.

We follow the recipe in [8, VII, 3.1] to prove that the space $\mathscr{F}_{\mu,\beta_1,\beta_2}$ is at most two dimensional. This follows in a rather straight forward manner. Therefore, we only provide an outline of the argument here. By definition, $\mathscr{F}_{\mu,\beta_1,\beta_2}$ is the space of functionals $\mathscr{D}\in\mathscr{E}'(H_{\beta_1}^\infty,H_{\beta_2}^\infty)$ satisfying

$$\begin{aligned}\mathscr{D}(n(t)\cdot\psi_1,n(t)\cdot\psi_2)&=\mathscr{D}(\psi_1,\psi_2)\\ \mathscr{D}(a(t)\cdot\psi_1,a(t)\cdot\psi_2)&=e^{(-1+\nu)t/2}\mathscr{D}(\psi_1,\psi_2).\end{aligned}$$

Using (2.1), these conditions translate to

$$\mathscr{D}(\psi_1(x-t),\psi_2(x-t)) = \mathscr{D}(\psi_1,\psi_2) \tag{3.13}$$

$$\mathscr{D}(\psi_1(e^{-t}x),\psi_2(e^{-t}x)) = e^{(1+\nu_1+\nu_2+\nu)t/2}\mathscr{D}(\psi_1,\psi_2). \tag{3.14}$$

By [8, VII, 3.1, (2)], condition (3.13) implies that $\mathscr{D}(\psi_1,\psi_2) = \mathscr{D}_0(\omega)$, where $\mathscr{D}_0$ is a one-dimensional distribution, and ω is the convolution

$$\omega(x) = \int \psi_1(x_1)\overline{\psi_2(x+x_1)}dx_1.$$

Condition (3.14) now implies that $\mathscr{D}_0(\omega) = e^{-(1+\nu_1+\nu_2+\nu)t/2}\mathscr{D}_0(\omega(e^{-t}x))$, or equivalently,

$$\mathscr{D}_0(\omega) = |\alpha|^{(1+\nu_1+\nu_2+\nu)/2}\mathscr{D}_0(\omega(\alpha x)).$$

This shows that $\mathscr{D}_0$ is a generalized homogeneous function of degree $(-1+\nu_1+\nu_2+\nu)/2$. The space of homogeneous functionals on $\mathbb{R}$ have been characterized completely and it is at most two dimensional. See [8, VII, 3.1, (6) and (7)] for more details.

Moreover, note that any functional in $\mathscr{E}_{\mu,\beta_1,\beta_2}$ can be realized as a functional in $\mathscr{F}_{\mu,\beta_1,\beta_2}$, via the equivalence of V_β with K_β for any $\beta\in\Omega_\Gamma$, and using the map ϕ in (2.3) between the Line model and the Kirillov model. This implies that the space $\mathscr{E}_{\mu,\beta_1,\beta_2}$ is also at most two dimensional. However, it can be easily checked that the functionals $B^1_{\mu,\beta_1,\beta_2}$ and $B^2_{\mu,\beta_1,\beta_2}$ defined by (3.11) and (3.12) are in $\mathscr{E}_{\mu,\beta_1,\beta_2}$. Furthermore, since the space of Kirillov model K_β, for any $\beta\in\Omega_\Gamma$, contains the space of smooth compactly supported functions $\mathrm{C}_c^\infty(\mathbb{R}\setminus\{0\})$, it can be easily deduced that $B^1_{\mu,\beta_1,\beta_2}$ and $B^2_{\mu,\beta_1,\beta_2}$ are linearly independent, thus proving the lemma. □

We start by proving the bound:

Lemma 3.5. *Let $0<\mu<1/4$. Then $B^1_{\mu,\beta_1,\beta_2}$ and $B^2_{\mu,\beta_1,\beta_2}$ belong to $\mathscr{E}_{\mu,\beta_1,\beta_2}$ and satisfy*

$$|B^1_{\mu,\beta_1,\beta_2}(f_{\beta_1},f_{\beta_2})| + |B^2_{\mu,\beta_1,\beta_2}(f_{\beta_1},f_{\beta_2})| \le S_{2,0}(X^{(1-\nu)/2}f_{\beta_1})S_{2,0}(X^{(1-\nu)/2}f_{\beta_2})\,.$$

Proof. Recall that for $j=1,2$, $X(\phi_j f) = ix\phi_j f$, then iX is a self-adjoint operator with spectrum $\mathbb{R}$. The spectral theorem then shows $X^s\phi_j f = e^{i\pi s/2}x^s\phi_j f$ for any $s\ge 0$. Therefore we have

$$\begin{aligned}
&|B^1_{\mu,\beta_1,\beta_2}(f_{\beta_1}\otimes f_{\beta_2})| + |B^2_{\mu,\beta_1,\beta_2}(f_{\beta_1},f_{\beta_2})|\\
&\le \int_{\mathbb{R}/\{0\}} |x|^{-(1+\nu)/2}|(\phi_1 f_{\beta_1})(x)\overline{(\phi_2 f_{\beta_2})(x)}|dx\\
&\le \left(\int_{\mathbb{R}/\{0\}} |x|^{-(1+\nu)/2}|(\phi_1 f_{\beta_1})(x)|^2dx\right)^{1/2}\left(\int_{\mathbb{R}/\{0\}} |x|^{-(1+\nu)/2}|(\phi_2 f_{\beta_2})(x)|^2dx\right)^{1/2}\\
&= \left(\int_{\mathbb{R}/\{0\}} |x^{(1-\nu)/2}(\phi_1 f_{\beta_1})(x)|^2\frac{dx}{|x|}\right)^{1/2}\left(\int_{\mathbb{R}/\{0\}} |x^{(1-\nu)/2}(\phi_2 f_{\beta_2})(x)|^2\frac{dx}{|x|}\right)^{1/2}\\
&= S_{2,0}(X^{(1-\nu)/2}\phi_1 f_{\beta_1})S_{2,0}(X^{(1-\nu)/2}\phi_2 f_{\beta_2})\,.
\end{aligned}$$

□

Since $\mathscr{D}^-_{\mu,\beta_1,\beta_2}$ belongs to $\mathscr{E}_{\mu,\beta_1,\beta_2}$, it can be written as a linear combination of $B^1_{\mu,\beta_1,\beta_2}$ and $B^2_{\mu,\beta_1,\beta_2}$ as follows.

Lemma 3.6. *There are constants $\mathscr{C}^1_{\mu,\beta_1,\beta_2}, \mathscr{C}^2_{\mu,\beta_1,\beta_2} \in \mathbb{C}$ such that*

$$|\mathscr{C}^1_{\mu,\beta_1,\beta_2}| + |\mathscr{C}^2_{\mu,\beta_1,\beta_2}| \ll (1+|\beta_1|)^{3/2}(1+|\beta_2|)^{3/2}$$

and

$$\mathscr{D}^-_{\mu,\beta_1,\beta_2} = \mathscr{C}^1_{\mu,\beta_1,\beta_2} B^1_{\mu,\beta_1,\beta_2} + \mathscr{C}^2_{\mu,\beta_1,\beta_2} B^2_{\mu,\beta_1,\beta_2}.$$

on $V^2_{\beta_1} \times V^2_{\beta_2}$.

Proof. Let $(f,g) \in (V_{\beta_1}, V_{\beta_2})$ be such that $\phi_1(f), \phi_2(g) \in C^\infty_c((1,2))$ are non-negative valued functions taking the value 1 on the interval $(5/4, 7/4)$. By a direct computation in the Kirillov model, we may further choose f, g such that for any integer $k \geq 0$,

$$S_{2,k}(f) \ll (1+|\beta_1|)^k$$
$$S_{2,k}(g) \ll (1+|\beta_2|)^k \,.$$

Using interpolation, the above bounds hold for any $k \in \mathbb{R}_+$. Since $(f,g) \in (V^\infty_{\beta_1}, V^\infty_{\beta_2})$ and $B^2_{\mu,\beta_1,\beta_2}$ vanishes on (f,g), we get

$$\mathscr{D}^-_\mu(f\overline{g}) = \mathscr{C}^1_{\mu,\beta_1,\beta_2} B^1_{\mu,\beta_1,\beta_2}(f,g).$$

Moreover,

$$B^1_{\mu,\beta_1,\beta_2}(f,g) = \int_1^2 |x|^{-(1+\nu)/2}(\phi_1 f)(x)\overline{(\phi_2 g)(x)}dx \gg 1.$$

Since the $W^{-3/2-\varepsilon}(M)$ norm of $\mathscr{D}^-_\mu$ is bounded by a constant $C>0$, using the fact that $W^s(M)$ is a Banach algebra for any $s > 3/2$, we get

$$\begin{aligned}|\mathscr{C}^1_{\mu,\beta_1,\beta_2}| \ll |\mathscr{D}^-_\mu(f\overline{g})| \ll S_{2,3/2+\varepsilon}(f\overline{g}) &\ll S_{2,3/2+\varepsilon}(f) S_{2,3/2+\varepsilon}(g) \\ &\ll (1+|\beta_1|)^{3/2+\varepsilon}(1+|\beta|)^{3/2+\varepsilon}\,,\end{aligned}$$

thus giving the bound on $\mathscr{C}^1_{\mu,\beta_1,\beta_2}$. A similar treatment implies the result for $\mathscr{C}^2_{\mu,\beta_1,\beta_2}$. □

We now use proposition 3.2 to obtain the following bound for $B^1_{\mu,\beta_1,\beta_2}$ and $B^2_{\mu,\beta_1,\beta_2}$.

Lemma 3.7. *For $\mu \in \Omega_\Gamma \cap (0,1/4)$ and $i = 1,2$, we have*

$$|B^i_{\mu,\beta_1,\beta_2}(f_{\beta_1} \star \sigma_H, f_{\beta_2} \star \sigma_H)| \ll (1+|a|^{-1})H^{-1} S_{2,(3-\nu)/2}(f_{\beta_1}) S_{2,(3-\nu)/2}(f_{\beta_2})\,.$$

$i = 1,2$.

Proof. We only deal with bounding $B^1_{\mu,\beta_1,\beta_2}$ here, the other case is analogous. We begin by applying lemma 3.5 to get

$$|B^1_{\mu,\beta_1,\beta_2}(f_{\beta_1} \star \sigma_H, f_{\beta_2} \star \sigma_H)| \ll S_{2,0}((X^{(1-\nu)/2}f)_{\beta_1} \star \sigma_H)) S_{2,0}((X^{(1-\nu)/2}f)_{\beta_2} \star \sigma_H)\,,$$

since X commutes with $\star\sigma_H$. We now invoke the estimate in proposition 3.2 to get

$$|B^1_{\mu,\beta_1,\beta_2}(f_{\beta_1} \star \sigma_H, f_{\beta_2} \star \sigma_H)| \ll (1+|a|^{-1})H^{-1} S_{2,(3-\nu)/2}(f_{\beta_1}) S_{2,(3-\nu)/2}(f_{\beta_2})\,.$$

□

Finally, we use these functionals to estimate $\mathscr{D}^-_\mu(|f \star \sigma_H|^2)$ and prove proposition 3.3.

Proof of Proposition 3.3. Using Lemma 3.7 and Lemma 3.6, we get

$$\begin{aligned}
&|\mathscr{D}_\mu^-(|f\star\sigma_H|^2)| \qquad (3.15)\\
&=\left|\mathscr{D}_\mu^-\left(\sum_{\beta_1,\beta_2\in\Omega_\Gamma}(f\star\sigma_H)_{\beta_1}(\overline{f\star\sigma_H})_{\beta_2}\right)\right|\le\sum_{\beta_1,\beta_2\in\Omega_\Gamma}\left|\mathscr{D}^-_{\mu,\beta_1,\beta_2}(f_{\beta_1}\star\sigma_H,f_{\beta_2}\star\sigma_H)\right|\\
&\ll(1+|a|^{-1})H^{-1}\sum_{\beta_1,\beta_2\in\Omega_\Gamma}(1+|\beta_1|)^{3/2+\varepsilon}S_{2,(3-\nu)/2}(f_{\beta_1})(1+|\beta_2|)^{3/2+\varepsilon}S_{2,(3-\nu)/2}(f_{\beta_2})\\
&\ll(1+|a|^{-1})H^{-1}\sum_{\beta_1,\beta_2\in\Omega_\Gamma}S_{2,(9-\nu)/2+2\varepsilon}(f_{\beta_1})S_{2,(9-\nu)/2+2\varepsilon}(f_{\beta_2}),
\end{aligned}$$

using the fact that $(1+|\beta_j|)^kS_{2,s}(f_{\beta_j})\ll S_{2,s+2k}(f_{\beta_j})$, for $j=1,2$, and for any $k,s\in\mathbb{R}_+$. Now using the Cauchy-Schwarz inequality and the Plancherel formula, we get

$$\begin{aligned}
&\ll\#\{\Omega_\Gamma\cap(0,1/4)\}(1+|a|^{-1})H^{-1}\left(\sum_{\beta\in\Omega_\Gamma}S_{2,(9-\nu)/2+2\varepsilon}(f_\beta)\right)^2\\
&\ll\#\{\Omega_\Gamma\cap(0,1/4)\}(1+|a|^{-1})H^{-1}\left(\sum_{\beta\in\Omega_\Gamma}(1+|\beta|)^{-1/2-\varepsilon/2}S_{2,(11-\nu)/2+3\varepsilon}(f_\beta)\right)^2\\
&\ll(1+|a|^{-1})H^{-1}\#\{\Omega_\Gamma\cap(0,1/4)\}(\sum_{\beta\in\Omega_\Gamma}(1+|\beta|)^{-1-\varepsilon})S_{2,(11-\nu)/2+3\varepsilon}(f)^2\,.
\end{aligned}$$

The Weyl's law for the distribution of eigenvalues [11, Lemma 2.28] implies that $\#\{\beta\in\Omega_\Gamma,|\beta|\le T_0\}\ll T_0$, thus implying $\sum_{\beta\in\Omega_\Gamma}|\beta|^{-1-\varepsilon}<\infty$. Further adding over all $\mu\in(0,1/4)\cap\Omega_\Gamma$ establishes the proposition. □

3.3. **Proof of theorem 1.2.**

We start by bounding the W^s norm of $|f\star\sigma_H|^2$.

Lemma 3.8. *Let $f\in \mathrm{C}^\infty(M)$, $H\ge1$ and $\varepsilon>0$. Let $\eta>0$ be such that for any $k\in\mathbb{N}$,*

$$S_{\infty,k}(f\star\sigma_H)\ll C_{a,k}H^{2k-\eta}S_{2,k+11/2+\varepsilon}(f),$$

where $C_{a,k}$ is a constant depending on a and k, satisfying $C_{a,k}\le C_{a,k+1}$. Then for any $s\in\mathbb{R}_+$, we have

$$S_{2,s}(|f\star\sigma_H|^2)\ll_\varepsilon(1+|a|^{-1/2})C_{a,\lceil s\rceil}H^{2s-1/2-\eta+\varepsilon}S_{2,s+11/2+\varepsilon}(f)^2,$$

where $\lceil s\rceil$ denotes the nearest integer greater than or equal to s.

Proof. Let $s\ge0$ be an integer. We can easily see that after using proposition 3.2 we get

$$S_{2,s}(|f\star\sigma_H|^2)^2\ll\sum_{j=0}^{s}S_{\infty,j}(f\star\sigma_H)^2S_{2,s-j}(f\star\sigma_H)^2\ll C_{a,s}(1+|a|)H^{4s-1-2\eta}S_{2,s+11/2+\varepsilon}(f)^4.$$

The bound can then be extended for any $s\in\mathbb{R}_+$, after using interpolation. □

Proof of theorem 1.2. We begin by observing that (3.5) implies that for any $x_0 \in M$, we have

$$|f \star \sigma_T(x_0)| \ll \sum_{j=0}^{s} \sum_{B \in \mathscr{O}_j} \nu_T(|Bf \star \sigma_H|^2)^{1/2} T^{2j} + HT^{2s-1} S_{2,s}(f). \tag{3.16}$$

We are now ready to finish the first part of theorem 1.2, namely the bound (1.5). Using (3.2), and Sobolev embedding, it is easy to establish the bound

$$S_{\infty,k}(f \star \sigma_H) \leq C_k H^{2k} S_{\infty,s}(f) \ll C'_k H^{2k} S_{2,s+3/2+\varepsilon}(f), \tag{3.17}$$

where C'_k is an increasing sequence. Thus, the hypothesis of lemma 3.8 is valid with $\eta = \eta_0 := 0$. This implies the bound

$$S_{2,s}(|f \star \sigma_H|^2) \ll_\varepsilon (1 + |a|^{-1}) C'_k H^{2s-1/2-\eta_0+\varepsilon} S_{2,s+11/2+\varepsilon}(f)^2.$$

Keeping the explicit dependence on η will be useful in proving the later part of the proof. By substituting the above bounds, along with the ones from proposition 3.2, and proposition 3.3 , into lemma 3.1, for any $\varepsilon > 0$, we have,

$$\nu_T(|Bf \star \sigma_H|^2) \ll_{\varepsilon,\Gamma} (1 + |a|^{-1}) \left(\frac{1}{H} + \frac{H^{1+\varepsilon}}{T^{1/2-\varepsilon}} + \frac{H^{7/2-\eta_0+\varepsilon}}{T} \right) (S_{2,11/2+\varepsilon}(Bf))^2,$$

Applying these bounds to (3.16), we get

$$S_{\infty,s}(f \star \sigma_T) \ll_{\varepsilon,\Gamma} (1 + |a|^{-1/2}) T^{2s} \left(H^{-1/2} + \frac{H^{7/4-\eta_0/2+\varepsilon}}{\sqrt{T}} \right) S_{2,s+11/2+\varepsilon}(f).$$

Optimizing, we set $H = T^{2/(9-2\eta_0)-\varepsilon}$, which yields

$$S_{\infty,s}(f \star \sigma_T) \ll_{\varepsilon,\Gamma} (1 + |a|^{-1/2}) T^{2s-1/(9-2\eta_0)+\varepsilon} S_{2,s+11/2+\varepsilon}(f), \tag{3.18}$$

thus proving the bound (1.5), upon recalling that $\eta_0 = 0$. The explicit dependence of ψ and T in (3.18) will be crucial in the application to the sparse equidistribution.

Henceforth, we assume that $a \neq 0$ is fixed. The validity of the bound (1.6) implies that the hypothesis of lemma 3.8 holds with $\eta = \eta_1 = 1/9 - \varepsilon$, and $C_{a,k} \ll_k (1 + |a|^{-1/2})$. Now, the process of obtaining (3.18) can be bootstrapped to obtain that for any $j \in \mathbb{N}$, we have

$$S_{\infty,s}(f \star \sigma_T) \ll_{\varepsilon,\Gamma,s,j} (1 + |a|^{-1/2})^{j+1} T^{2s-1/(9-2\eta_j)+\varepsilon} S_{2,s+11/2+\varepsilon}(f), \tag{3.19}$$

where $\eta_0 = 0$, and the sequence η_j satisfying $\eta_{j+1} := \frac{1}{9-2\eta_j}$. It can be easily seen that the sequence η_j converges to $(9 - \sqrt{73})/4$, which is a solution to the quadratic equation $2y^2 - 9y + 1$, thus proving (1.6), and the theorem. $\square$

4. Proof of theorem 1.3 and Theorem 1.4

4.1. **Proof of theorem 1.3.** We follow a variant of the recipe in the proof of [18, Theorem 3.1]. We start by proving effective equidistribution for arithmetic progressions:

Lemma 4.1. *Let $x_0 \in M$, and let $f \in \mathrm{C}^\infty(M)$ be such that $\int_M f(g)dg = 0$. Let b, b_1, ε be positive numbers satisfying $b + \varepsilon < \frac{(1-2\alpha_0)^2}{8(3-2\alpha_0)}$ and $b_1 + \varepsilon < 1/9$. Then*

$$\left| \sum_{1 \le j \le K^{r-1}} f(x_0 n(Kj)) \right| \ll_{f,\varepsilon} K^{r-1-\varepsilon},$$

for any $r \ge 3/(b + 2b_1)$.

Proof. Let $g_\delta = \max(\delta^{-2}(\delta - |t|), 0)$ be a function on $\mathbb{R}$. Using the Poisson summation formula, it can be easily seen that

$$\sum_{j \in \mathbb{Z}} g_\delta(t + jK) = \sum_{k \in \mathbb{Z}} \exp(2\pi i K^{-1} kt) a_k,$$

where $a_\lambda = K^{-1} \int_{\mathbb{R}} \exp(-2\pi i \lambda t) g_\delta(t) dt$. It can be easily seen that $|a_k| \le K^{-1}$ and $\sum_k |a_k| \ll \delta^{-1}$. Note that

$$\int_0^T \left(\sum_{j \in \mathbb{Z}} g_\delta(t + Kj) \right) f(x_0 n(t)) dt = \sum_{k \in \mathbb{Z}} a_k \int_0^T \exp(2\pi i K^{-1} kt) f(x_0 n(t)) dt. \tag{4.1}$$

We use the bound in lemma 1.1 to bound the integrals on the right hand side of (4.1) when $|k| < K_0$, say, and use the bound of theorem 1.2 to bound the rest. We then have

$$\left| \int_0^T \left(\sum_{j \in \mathbb{Z}} g_\delta(t + Kj) \right) f(x_0 n(t)) dt \right| \ll_{\Gamma,f,\varepsilon} K_0 K^{-1} T^{1-b-\varepsilon} + (K/K_0)^{1/2} T^{1-b_1-\varepsilon} \delta^{-1}. \tag{4.2}$$

Moreover, since g_δ is supported in a δ neighborhood of 0, and it has integral 1, we can easily deduce that

$$\left| \int_0^T \left(\sum_{j \in \mathbb{Z}} g_\delta(t + Kj) \right) f(x_0 n(t)) dt - \sum_{\substack{j \in \mathbb{Z} \\ 1 \le Kj \le T}} f(x_0 n(Kj)) \right| \ll_f (1 + TK^{-1}\delta).$$

Combining with (4.2), we get

$$\left| \sum_{\substack{j \in \mathbb{Z} \\ 1 \le Kj \le T}} f(x_0 n(Kj)) \right| \ll_{\Gamma,f,\varepsilon} 1 + TK^{-1}\delta + K_0 K^{-1} T^{1-b-\varepsilon} + (K/K_0)^{1/2} T^{1-b_1-\varepsilon} \delta^{-1}. \tag{4.3}$$

Choose $T = K^r$, $K_0 = K^{rb}$, and $\delta = K^{-\varepsilon}$ to get

$$\left| \sum_{1 \le j \le K^{r-1}} f(x_0 n(Kj)) \right| \ll_{\Gamma,f,\varepsilon} 1 + K^{r-1-\varepsilon} + K^{r-1-r\varepsilon} + K^{r-1+(3/2-r(b/2+b_1))-(r-1)\varepsilon}.$$

The lemma now follows upon choosing $r \ge 3/(b + 2b_1)$.

∎

Proof of theorem 1.3. By possibly subtracting f by a constant, we may assume that f is a function of zero average. The implied constants appearing in the proof here are independent of N, but may depend on f, Γ, and a parameter ε chosen in due course.

The theorem follows easily from lemma 4.1 after approximating the sequence $\{j^{1+\gamma} : 0 \leq j \leq N\}$ by a union of arithmetic progressions. In particular, let $N_0 \in \mathbb{N}$ be a large number. For a small t we have

$$(N_0 + t)^{1+\gamma} = N_0^{1+\gamma}(1 + t/N_0)^{1+\gamma} = N_0^{1+\gamma} + (1 + \gamma)tN_0^{\gamma} + O(t^2 N_0^{\gamma-1}). \tag{4.4}$$

This is a good approximation for $t \ll N_0^{(1-\gamma)/2-\varepsilon}$, where ε is a small positive number. For $N_1 = N^{1-\varepsilon}$, we write $\{j^{\gamma+1} : 0 \leq j \leq N\} = \{j^{\gamma+1} : 0 \leq j \leq N_1 - 1\} \cup \{j^{\gamma+1} : N_1 \leq j \leq N\}$. The second set can be decomposed into a disjoint union of $L - 1$ sets of the form

$$\cup_{k=1}^{L-1} \{N_k^{1+\gamma}, (N_k + 1)^{1+\gamma}, ..., (N_k + N_k^{(1-\gamma)/2-\varepsilon})^{1+\gamma}\} \cup \{N_L^{1+\gamma}, ..., N^{1+\gamma}\},$$

where for any $k \geq 1$, $N_{k+1} = N_k + N_k^{(1-\gamma)/2-\varepsilon} + 1$, and $N_L \leq N \leq N_{L+1}$. Clearly, the terms in the tail $N^{-1} \sum_{j=N_L}^{N} f(x_0 n(j^{1+\gamma}))$ can be bound appropriately. Since each of these above sets have $N_k^{(1-\gamma)/2-\varepsilon}$ elements, we have

$$\sum_{k=1}^{L-1} N_k^{(1-\gamma)/2-\varepsilon} \ll N.$$

For each $1 \leq k < L$, we can now apply the lemma 4.1 along with (4.4) as long as $N_k^{(1-\gamma)/2-\varepsilon} > N_k^{\gamma(r_0-1)}$, where $r_0 = 3/(b + 2b_1)$. If $\frac{1-\gamma}{2\gamma} > r_0 - 1$, equivalently if $\gamma < 1/(2r_0 - 1) = 1/(6/(b+2b_1) - 1)$, a suitable value of $\varepsilon > 0$ can be chosen so that the above condition as well as the hypothesis of lemma 4.1 hold. The theorem now follows easily from the following estimates

$$\begin{aligned}\frac{1}{N}\left|\sum_{k=1}^{L-1} \sum_{j=1}^{N_k^{(1-\gamma)/2-\varepsilon}} f(x_0 n((N_k + j)^{1+\gamma}))\right| &\ll \frac{1}{N}\sum_{k=1}^{L-1}\left|\sum_{j=1}^{N_k^{(1-\gamma)/2-\varepsilon}} f(x_0 n(N_k + N_k^{\gamma} j))\right| \\ &\ll \frac{1}{N}\sum_{k=1}^{L-1} N_k^{(1-\gamma)/2-\varepsilon-\gamma\varepsilon} \ll N_1^{-\gamma\varepsilon} \ll N^{-(1-\varepsilon)\gamma\varepsilon}.\end{aligned}$$

$\square$

4.2. **Proof of theorem 1.4.** Let $\mathrm{L}_0^2(M) := \left\{f \in L^2(M) : \int_M f(g)\, d_\rho g = 0\right\}$, where $\rho \in W^6(M)$ is a positive function. Using the commutation relations

$$[X_\rho, Y] = (\frac{Y\rho}{\rho} - 1)X_\rho\,,\ [X_\rho, Z] = \frac{Y}{\rho} + \frac{Z\rho}{\rho}X_\rho\,,$$

and solving a system of O.D.E's, the tangent flow $\{Dn_t^\rho\}$ on TM is computed in [7, Lemma 1]. It follows that there are continuous functions $y_{X_\rho}, z_{X_\rho}, z_Y : \mathbb{R} \to \mathbb{R}$ satisfying

$$|y_{X_\rho}(t)| + |z_Y(t)| \ll_\rho |t|\,,\ |z_{X_\rho}(t)| \ll_\rho |t|^2\,,$$

such that for every $f \in C^\infty(M)$,

$$\begin{aligned} Y(n^\rho(t)\cdot f) &= \big(y_{X_\rho}(t)n^\rho(t)\cdot X_\rho + n^\rho(t)\cdot Y\big)(f)\,,\\ Z(n^\rho(t)\cdot f) &= \big(z_{X_\rho}(t)n^\rho(t)\cdot X_\rho + z_Y(t)n^\rho(t)\cdot Y + n^\rho(t)\cdot Z\big)(f)\,,\\ X_\rho(n^\rho(t)\cdot f) &= n^\rho(t)\cdot X_\rho f\,. \end{aligned} \tag{4.5}$$

The main step in proving theorem 1.4 is the following lemma.

Lemma 4.2. *For all $f \in C^\infty(M)\cap \mathrm{L}^2_0(M)$, and all $T \geq 1$, and $\varepsilon > 0$,*

$$S_{\infty,0}(f\star\sigma^\rho_T) \ll_\rho T^{-(1-\alpha_0)^2/(100-4\alpha_0)+\varepsilon} S_{2,15/2}(f)\,.,$$

where

$$f\star\sigma^\rho_T(f)(x) := \frac{1}{T}\int_0^T \psi(t) f(xn^\rho(t))\,dt\,.$$

Proof. The proof will follow by combining the argument in [18, Lemma 3.1] with results on the quantitative equidistribution and quantitative mixing of $\{n^\rho(t)\}_{t\in\mathbb{R}}$ in [7, Theorems 2,3].

We provide slightly weaker versions of these results here. Recall that $\alpha_0 = \sqrt{1-4\lambda_1} \in (0,1]$, where λ_1 is the spectral gap of the Laplace-Beltrami operator on M.

Lemma 4.3 (Theorems 2 and 3 in [7])**.**

- *For any $r > 3$, $T \geq 1$, $x_0 \in M$, and $f \in W^r(M)$, we have*

$$\Big|\int_0^T f(x_0 n^\rho(t))\,dt - \int_M f(g)\,d_\rho g\Big| \ll_{r,\rho} T^{-\frac{1-\alpha_0}{2}}(1+\log T) S_{2,r}(f)$$

- *For any $r > 11/2$, $(x,t) \in M\times\mathbb{R}_{\geq 1}$, and for any $f \in W^r(M)\cap\mathrm{L}^2_0(M)$ and any $g \in W^r(M)$,*

$$|\langle n^\rho(t) f, g\rangle_{L^2(M,\mathrm{vol}_\rho)}| \ll_{r,\rho} S_{2,r}(f)S_{2,r}(g) t^{-\frac{1-\alpha_0}{2}}(1+\log t)\,.$$

Let $\nu^\rho_T(f) := \frac{1}{T}\int_0^T f(x_0 n^\rho(t))\,dt$. We can easily derive a bound analogous to (3.5), and [18, (3.5)]:

$$\begin{aligned} |S_{\infty,0}(f\star\sigma^\rho_T)| &\ll \frac{H}{T}S_{\infty,0}(f) + \sqrt{\nu^\rho_T(|f\star\sigma^\rho_H|^2)}\\ &= \frac{H}{T}S_{\infty,0}(f) + \left(\frac{1}{H^2}\int_{(h_1,h_2)\in[0,H]^2} |\nu_T(n^\rho(h_1)f\cdot\overline{n^\rho(h_2)f})|\,dh_1dh_2\right)^{1/2}\\ &\ll_\rho \frac{H}{T}S_{\infty,0}(f) + \left(\frac{1}{H^2}\int_{(h_1,h_2)\in[0,H]^2} |\langle n^\rho(h_1-h_2)f, f\rangle|\,dh_1dh_2\right)^{1/2}\\ &\qquad + \left(T^{-\frac{1-\alpha_0}{2}}(1+\log T)\sup_{(h_1,h_2)\in[0,H]^2} S_{2,6}(n^\rho(h_1)f\cdot\overline{n^\rho(h_2)f})\right)^{1/2}. \end{aligned} \tag{4.6}$$

Using (4.5), we also can derive bounds analogous to (3.17)

$$S_{2,6}(n^\rho(h_1)\cdot f\ \overline{n^\rho(h_2)\cdot f}) \ll_\rho (1+|h_1|+|h_2|)^{12} S_{2,15/2}(f)^2\,.$$

Applying the Sobolev embedding theorem, along with the mixing statement in lemma 4.3, to (4.6) gives that for all $\varepsilon > 0$, we have

$$(4.6) \ll_\rho (\frac{H}{T} + H^{(\alpha_0-1)/4+\varepsilon} + \frac{H^6}{T^{\frac{1-\alpha_0}{4}-\varepsilon}}) S_{2,15/2}(f). \tag{4.7}$$

Optimizing, set $H = T^{(1-\alpha_0)/(25-\alpha_0)-\varepsilon}$, and we get

$$(4.7) \ll_\rho T^{-(1-\alpha_0)^2/(100-4\alpha_0)+\varepsilon} S_{2,15/2}(f)\,.$$

This concludes the proof of lemma 4.2. □

Proof of Theorem 1.4. Let $f \in L^2_0(M)$. The method used to prove lemma 4.1 can be recycled here upon setting $K_0 = 1$, $b = b_1 = (1-\alpha_0)^2/(100-4\alpha_0)$, and replacing $n(t)$ with $n^\rho(t)$. A formula analogous to (4.3) then gives

$$\left| \sum_{\substack{j\in\mathbb{Z} \\ 1\le j\le T}} f(x_0 n^\rho(Kj)) \right| \ll_\rho (1 + TK^{-1}\delta + T^{1-b-\varepsilon}\delta^{-1}) S_{2,15/2}(f)\,.$$

As before, let $T = K^r$ and $\delta = K^{-\varepsilon}$ to get

$$\left| \sum_{\substack{j\in\mathbb{Z} \\ 1\le j\le K^r}} f(x_0 n^\rho(Kj)) \right| \ll_\rho (1 + K^{r-1+\varepsilon} + K^{r(1-b-\varepsilon)-\varepsilon}) S_{2,15/2}(f)\,. \tag{4.8}$$

Now using the notation in the proof of theorem 1.3, we approximate the sequence $\{j^{1+\gamma} : 0 \le j \le N\}$ by a union of arithmetic progressions. We use (4.8) in the place of lemma 4.1, and conclude that theorem 1.4 holds for any $\gamma < \frac{b}{2}$. □

References

[1] R. A. Adams , *Sobolev Spaces*, Pure Appl. Math. 65, Academic Press, New York, 1975.

[2] E. M. Baruch. *The classical Hankel transform in the Kirillov model of the discrete series*, Integral Transforms Spec. Funct. 24 (2013), no. 5, 339356.

[3] J. Bernstein, A. Reznikov. *Sobolev norms of automorphic functionals and Fourier coefficients of cusp forms*, C. R. Acad. Sci. Paris S er. I Math., 327(2):111116, 1998.

[4] L. Flaminio, G. Forni, *Invariant distributions and time averages for horocycle flows*, Duke math J, Volume 119, Number 3 (2003), 465-526.

[5] L. Flaminio, G. Forni, *On effective equidistribution for higher step nilflows*, arXiv:1407.3640 .

[6] L. Flaminio, G. Forni, J. Tanis, *Twisted Horocycle flow*, in preparation.

[7] G. Forni, C. Ulcigrai, *Time-Changes of Horocycle Flows*, Journal of Modern Dynamics, Vol. 6, no. 2, 2012, 251- 273

[8] I M Gel'fand , M I Graev , N Ya Vilenkin, E. Saletan, *Generalized Functions - Vol 5: Integral Geometry and Representation theory.*

[9] D. Goldfeld, J. Hundley, *Automorphic Representations and L-Functions for the General Linear Group: Volume 1* , Cambridge Studies in Advanced Mathematics (Book 129), Cambridge University Press; 1 edition (May 30, 2011)

[10] A. Good. *Cusp forms and eigenfunctions of the Laplacian*, Math. Ann, 255: 523438, 1981.

[11] D. Hejhal, *The Selberg Trace Formula for $PSL(2,\mathbb{R})$*, Lecture Notes in Mathematics, Volume 1001 1983.

[12] E. Nelson, *Analytic Vectors*, Annals of Mathematics, Second Series, Vol. 70, No. 3, 572-615

[13] J. Lions *Non-homogeneous boundary value problems and applications. Vol 1.* Springer-Verlag New York, New York, 1972.

[14] G. A. Margulis. *Problems and conjectures in rigidity theory*, Mathematics: frontiers and perspectives, (2000) 161-174, Amer. Math. Soc., Providence, RI.

[15] J. Repka. *Tensor products of unitary representations of* $SL_2(\widehat{\mathbb{R}}^2)$. Amer. J. Math. 100 (1978), no. 4, 836-844.

[16] P. Sarnak, A. Ubis *The horocycle flow at prime times*. 2011, arXiv:1110.0777.

[17] J. Tanis, The *Cohomological equation and Invariant Distributions for Horocycle Maps*, Ergodic Theory & Dynamical Systems. (2012) CJO 2012 doi:10.1017/etds.2012.125

[18] A. Venkatesh, *Sparse equidistribution problems, period bounds and subconvexity* , Annals of Mathematics, Volume 172 (2010), Issue 2, pages 989-1094.